

基于生理大数据的情绪识别研究进展

赵国朕¹ 宋金晶¹ 葛 燕¹ 刘永进² 姚 林³ 文 涛³

¹(中国科学院心理研究所行为科学院重点实验室 北京 100101)

²(清华信息科学与技术国家实验室(筹) 北京 100084)

³(中国移动研究院 北京 100055)

(liuyongjin@tsinghua.edu.cn)

Advances in Emotion Recognition Based on Physiological Big Data

Zhao Guozhen¹, Song Jinjing¹, Ge Yan¹, Liu Yongjin², Yao Lin³, and Wen Tao³

¹(Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101)

²(Tsinghua National Laboratory for Information Science and Technology, Beijing 100084)

³(China Mobile Research Institute, Beijing 100055)

Abstract Affective computing (AC) is a new field of emotion research along with the development of computing technology and human-machine interaction technology. Emotion recognition is a crucial part of the AC research framework. Emotion recognition based on physiological signals provides richer information without deception than other techniques such as facial expression, tone of voice, and gestures. Many studies of emotion recognition have been conducted, but the classification accuracy is diverse due to variability in stimuli, emotion categories, devices, feature extraction and machine learning algorithms. This paper reviews all works that cited DEAP dataset (a public available dataset which uses music video to induce emotion and record EEG and peripheral physiological signals) and introduces detailed methods and algorithms on feature extraction, normalization, dimension reduction, emotion classification, and cross validation. Eventually, this work presents the application of AC on game development, multimedia production, interactive experience, and social network as well as the current limitations and the direction of future investigation.

Key words emotion recognition; electroencephalograph (EEG); peripheral physiological signal; feature extraction; machine learning

摘 要 随着计算技术和人机交互技术的不断发展,情感计算(affective computing, AC)逐渐成为情绪研究的新兴领域,而情绪识别(emotion recognition)又是情感计算中不可或缺的一环.基于生理信号的情绪识别方法比其他指标如面部表情、语音语调、身体姿势等更难以伪装,也能提供更丰富的信息.目前基于生理信号的情绪识别研究很多,但受到各种因素的影响,如刺激选取、诱发情绪的类别、采集设备、特征提取方法、不同的降维和分类算法等,各个研究的识别准确率差异性很大,很难进行比较.针对使用DEAP数据库(用音乐视频诱发情绪并采集脑电及外周生理信号的公开数据库)进行情绪识别的16篇文章做了梳理;对特征提取、数据标准化、降维、情绪分类、交叉检验等方法做了详细的解释和比较;最后分析了现阶段情绪识别在游戏开发、多媒体制作、交互体验、社交网络中的初步探索和应用,以及情绪识别和情感计算目前存在的问题及未来发展的方向.

关键词 情绪识别;脑电;外周生理信号;特征提取;机器学习

中图法分类号 TP18

收稿日期:2015-07-06;修回日期:2015-10-16

基金项目:国家自然科学基金项目(31300852,61322206)

This work was supported by the National Natural Science Foundation of China (31300852,61322206).

情绪是伴随着认知和意识过程产生的心理和生理状态,在人类交流中扮演着非常重要的角色。对情绪的研究由来已久,近年来,随着感知技术和人机交互技术的发展,情感计算(affective computing, AC)逐渐成为情绪研究的新兴领域。关于人机情感交互的思想由 Minsky^[1]最先提出,其在《Society of Mind》一书中指出“问题的关键不是智能机器是否需要情绪,而在于没有情绪的机器能否实现智能”。后来 Picard^[2]将“与情感有关、由情感引发或者能够影响情感因素的计算”界定为情感计算。

具体而言,情感计算是通过构建情绪状态而建立的计算模型,它对行为和生理信号进行分析,并基于测量到的情绪状态在人机之间建立情绪交互。其最重要的一个环节是“情绪识别(emotion recognition)”,即通过用户的行为和生理反应来预测估计相应的情绪状态^[3]。

基于脑电(electroencephalograph, EEG)的情绪识别^[4]和基于生理信号的情绪识别^[5]已经有研究者进行过综述。这2篇综述主要是从情绪的生理基础、情绪分类模型、诱发方法、脑电和生理信号的预处理、特征提取、降维和识别等方面进行了整理和介绍,但综述内容并不系统,涉及的文献也不完整。此外,考虑到情绪识别研究中的差异性(如刺激选取方法、诱发情绪的类别、采集设备、特征提取、降维和分类算法),情绪识别的准确率变化很大,难以进行比较。因此,本文选取了一个用音乐视频诱发情绪并采集脑电和外周生理信号的公开数据库 DEAP^[6],将使用该数据库对情绪识别的研究做了细致的梳理,从特征提取、数据标准化、特征降维、分类器、交叉检验(cross validation, CV)5个方面进行分析和比较,探讨各因素在情绪识别研究中的应用效果以及各因素中不同算法间的优缺点。

本文中的生理大数据是指面向情绪识别的脑电及外周生理(皮肤电、心电、呼吸、皮肤温度、肌电、眼电)信号。尽管在累积数据的规模上还不到互联网、社交媒体、医疗健康等领域超大数据的 PB 甚至 EB 级别,但其数据仍具有规模大(volume)、模态多样(variety)、速度快(velocity)和真实性(veracity)的“4V”特点^[7]。以 DEAP 数据库为例,32 人共计 1 280 min 的生理数据大小为 5.8 GB,按照单个医疗机构每日 200 人次的观测推算,数据规模已达到 TB 级别,其 40 通道的生理信号可提取的时域、频域、时频域特征从几百到上千不等。基于 DEAP 的实时情绪识别系统应在其刺激材料的持续时间(1 min)甚

至更短时间内对情绪状态作出判断。考虑到生理信号的时变、不稳定、噪声强等特点,在短时间内从大规模生理信号中提取干净、真实、可信的特征并以此训练、建立分类模型并预测情绪状态是非常具有挑战性的研究内容。此外,本文重点强调的是基于个体用户的生理信号预测其情绪状态,不同于基于语音或者网络大数据的(例如用户在微博等社交媒体上发布的带有感情色彩的文字、照片和视频)面向群体用户情绪状态预测的技术。生理信号的时域、频域和时频域特征有着较强的心理学和生理学基础,不需要人工规则构造特征;依赖于利用标准化的刺激材料诱发单一的目标情绪;利用用户的主观评价标记(分类)训练样本并采用有监督的机器学习算法;面向人机间情绪交互的实时、高效的情绪识别算法。

1 情绪的理论模型

作为“情绪识别”的对象,首先需要明确“情绪”的含义。Izard^[8]在一项通过对 34 名科学家的调研中提炼出情绪的构成:神经回路、反应系统以及一种激发和综合了认知及行为的感受状态。这可视为近 30 余年研究者对情绪的认知集合;然而,这种关于情绪的理论对“以实验技术操作情绪”和“以工程手段提取和判别情绪特征”而言缺乏可操作性。因此在研究情绪识别时,研究者通常采用以下 2 种视角来建构和理解情绪空间:离散模型(discrete model)和维度模型(dimensional model)。

1) 离散模型认为情绪空间由数量有限的离散的基本情绪构成。尽管不同研究者对基本情绪的认识并不一致,但多数研究倾向于认为至少存在以下 6 种基本情绪:高兴、悲伤、惊讶、恐惧、愤怒和厌恶^[9]。2) 维度模型则认为情绪空间可以由二维的效价-唤醒度(valence-arousal, VA)或是三维的愉悦度-唤醒度-优势度(pleasure-arousal-dominance, PAD)的坐标体系进行表达^[10]。其中,愉悦度是指个体感受到的愉悦程度,即正负特性(效价);唤醒度是指个体感受到情绪的心理警觉水平和生理激活的强烈程度;优势度是指个体处于控制还是被控制的地位^[11]。目前,基于 DEAP 数据库的情绪识别研究绝大多数建立在情绪维度模型的基础上。

以“离散”视角建立的情绪模型,其提炼出的基本情绪具有较好的普遍辨识性和自然属性,在神经系统层面也表现出了相对特异的反应模式^[9,12-13]。而以“维度”视角建立的情绪模型,即使在难以对某种特定

感受进行命名的情况下,该模型也可以将不同情绪用数值在坐标空间中进行标记^[14].从操作层面分析,尽管维度模型显然比离散模型更具操作性,但是当维度模型所要识别的情绪在效价和唤醒度上相近时,该方法对情绪的区分效力会相应下降^[15].这是由于在度量维度上相近的情绪(如焦虑和愤怒)可能具有各自独特的表象,并诱发完全不同的行为反应^[16].

近年来,研究人员们提出一种融合 2 种情绪模型的观点:将具有显著特征的离散的情绪状态以维度化的方式表达,可在一定程度上实现 2 种模型的相互融合,如图 1 所示.例如,愤怒可表示为具有消极效价的、较高唤醒的情绪^[17];恐惧则可表征为一种消极的、更高唤醒的情绪^[18].此外,同一种情绪也会因为刺激的不同而使感受到的情绪强烈程度不同.看到婴儿微笑会特别高兴,而看到美丽的日落可能高兴程度没有那么强烈^[18].从“情绪识别”的角度出发,情绪的维度模型更适合从量化的角度操纵和定义情绪;同时通过对生理信号的分析,也可以在一定程度上弥补“效价-唤醒度”模型在二维空间上缺乏分辨力的不足.

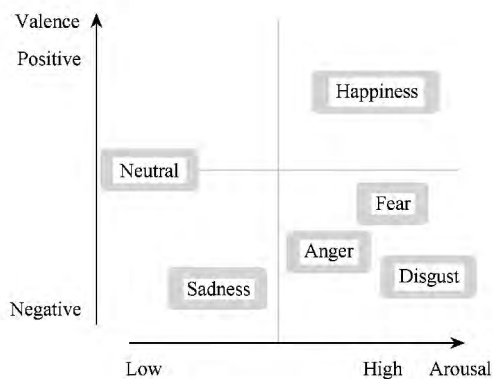


Fig. 1 Representing discrete emotions within a dimensional framework^[18].

图 1 将离散情绪表征到维度框架上

2 情绪诱发

情绪诱发是情绪识别研究中的重要一环.常用的情绪诱发方法包括自我诱发(如回忆^[19-21])、情景诱发(如玩游戏^[22-23])和事件诱发.其中以事件诱发的研究最多,即通过图片、声音、视频或任何可以唤起情绪的材料来诱发被试的情绪.

2.1 视觉刺激

视觉刺激是最常见的诱发情绪的材料,通过给被试呈现情绪性的图片、面部表情、文字等来诱发被

试的目标情绪是情绪诱发最常用的方法.经过标准化的视觉刺激材料,如国际情感图片库(international affective picture system, IAPS)^[24-25]、面孔表情图片库(pictures of facial affect, POFA)^[26]、英语情感词库(affective norms for English words, ANEW)^[27]等,均已在情绪识别领域中得到广泛的引用.虽然研究表明基本情绪具有跨文化的一致性^[28],但由于文化背景、社会学习等差异,很多国家都针对国际标准库进行相应的修订和完善.国内研究者也针对国际标准材料库进行了本土化修订,建立了适用于中国的情绪刺激材料,如中国情感图片库(Chinese affective picture system, CAPS)^[29]等.

2.2 听觉刺激

用听觉刺激诱发情绪比视觉刺激更为深入和持久,但目前经过标准化评定的听觉刺激材料很少^[30].现有的研究中,大多数是研究者自己制作的听觉刺激材料,包括电影背景音乐、歌曲、声音等,也有经过标准化的听觉刺激材料库,如国际情感数码声音库(international affective digitized sounds, IADS)^[31].国内研究者针对 IADS 编制了适用于中国被试的听觉刺激材料,即中国情感数码声音库(Chinese affective digital sounds, CADS)^[32].

2.3 视听觉刺激

早期的情绪识别研究主要使用图片或声音作为诱发情绪的材料.近年来,越来越多的情绪识别研究开始使用视听刺激结合的方式来诱发情绪.常用的视听觉刺激包括电影片段、音乐视频,或是呈现情绪图片的同时播放相应情绪的音乐^[33].相比于单纯的视、听觉刺激,电影与实际生活中的情境更相似,内容和情节更为丰富,具有动态性、不可欺骗性和高生态效度^[34].当然,电影诱发的情绪比较复杂,识别难度更大.常用的电影库包括 Gross and Levenson^[34]建立的美国情绪电影库、Schaefer 等人^[35]建立的法国情绪电影库、罗跃嘉等人^[36]建立的中国情感电影库和梁育琦等人^[37]建立的基于台湾被试的华语情绪电影库等.这些电影库都是基于离散情绪进行编制的.

2.4 DEAP 数据库

和电影比起来,音乐视频能够表达的情绪更为单一、连贯,更容易进入情绪状态,不像电影片段那样需要慢慢看情节才能诱发出情绪. Koelstra, Muhl 等人^[6]在情绪维度模型的基础上选取了 40 段音乐视频作为刺激材料,并采集了 32 名被试在观看 40 段音乐视频时的生理信号和脸部录像.生理信号共 40 导,其中脑电信号 32 导,外周生理信号(皮肤电、血压、呼吸、皮肤温度、肌电、眼电)8 导,组成了 DEAP

数据库,且为公开共享资源. DEAP 数据库还包含每名被试在观看每段视频后在效价、唤醒度、优势度和喜欢度 4 个维度上的主观评分,并以此对生理数据进行标记(有监督学习,详见 3.4 节). 截止投稿前,一共有 16 篇发表在期刊或会议上的文章使用

DEAP 数据库中的生理信号做情绪识别,如表 1 所示. 其中的 4 篇文章使用 DEAP 数据库建立了情绪识别系统,但没有详细报告与情绪识别相关的特征提取、分类算法、准确率等内容,因此不在表 1 中体现. 表 1 中出现的缩写词见第 3 节.

Table 1 Summary of Articles that Use Physiological Signals of DEAP Dataset for Emotion Recognition

表 1 使用 DEAP 数据库生理信号做情绪识别的文章汇总

Reference	Emotion/ Classes	EEG Features (#) ¹⁾	Peripheral Physiological Features(#) ¹⁾	Feature Extraction	Dimension Reduction	Window Length	Classifier	EEG Acc ²⁾ / % /Classes	Peripheral Acc ²⁾ / % /Classes	Cross Validation
Ref[6]	Arousal/2 Valence/2 Liking/2	PSD (160) ASM (56)	GSR(20) ECG(13) RSP(19) SKT(4) EOG(4)	Welch's Method	LDA		NB	62.0/2 57.6/2 55.4/2	57.0/2 62.7/2 59.1/2	SD
Ref[38]	Arousal/2 Valence/2 Liking/2 Dominance/2	PSD(384) ASM(168)		DT-CWPT	SVD- QRcp and F-Ratio		SVM	64.3/2 66.2/2 70.2/2 68.9/2		SD
Ref[39]	Arousal/2 Valence/2 Liking/2 Dominance/2	PSD(384) ASM(168)		DT-CWPT	SVD- QRcp and F-Ratio		SVM	65.3/2 66.9/2 71.2/2 69.1/2		SD
Ref[40]	Arousal/2 Valence/2 Liking/2 Dominance/2	PSD(160) ASM(70)		Generative Model	PCA	1s/2s/4s/8s with 1s step size	SVM KNN NB	68.4/2 76.9/2 75.3/2 73.9/2		Single-trial SD
Ref[41]	Arousal/2 Valence/2 Liking/2 Dominance/2	PSD(160) ASM(70)		Generative Model	PCA	1s with 0.1s step size	SVR	68.4/2 76.9/2 75.3/2 73.9/2		Single-trial SD
Ref[42]	Arousal/2 Arousal/3 Valence/2 Valence/3	PSD(160) ASM(56)	GSR(20) ECG(13) RSP(13) SKT(4) EOG(4)		Recursive Feature Elimination		SVM	85.71/2 ³⁾ 63.03/3 ³⁾ 75.49/2 ³⁾ 53.26/3 ³⁾	85.71/2 ³⁾ 63.03/3 ³⁾ 75.49/2 ³⁾ 53.26/3 ³⁾	
Ref[43]	Arousal/2 Arousal/3 Valence/2 Valence/3	PSD(128) ASM(244)		FFT	Pearson Correlation	2s with 1s step size	Bayes with a Perceptron Convergence Algorithm	70.1/2 55.2/3 70.9/2 55.4/3		
Ref[44]	Arousal/2 Arousal/3 Valence/2 Valence/3	Autore- gressive Coefficient		Autore- gressive Model	SFS Davies- Bouldin Index	1s	KNN QDA LDA	74.2/2 65.16/3 72.33/2 61.10/3		SID
Ref[45]	High Arousal Valence/2 Low Valence Arousal/2	Sample Entropy at Beta Band		WPT	K-S Test		SVM	80.43/2 71.16/2		3-fold CV SID
Ref[46]	Arousal/3 Valence/3	PSD(160) ASM(70)		FFT	PCA		DLN SVM NB	46.03/3 49.52/3		SID
Ref[47]	Valence/5	Statistical Features at Alpha Band			LDA		SVM FLEC ⁴⁾ NB	63.13/5		10-fold CV
Ref[48]	Discrete Model	PSD(12)	Spectrum Entropy	DWT	LDA		SVM MLP MMC ⁵⁾	81.45 (Average)	78.86 (Average)	10-fold CV

Notes: 1) # stands for the number of features; 2) Acc stands for accuracy; 3) the data is obtained by using the features extracted from both EEG and peripheral signals; 4) FLEC stands for fuzzy logic based emotion classification algorithm; 5) MMC stands for meta-multiclass.

3 基于生理大数据的情绪识别

情绪识别研究一般包括情绪诱发、生理信号采集、数据预处理、特征提取、特征降维、情绪学习和分

类6个步骤,如图2所示.由于DEAP数据库在情绪诱发、生理信号采集和数据预处理3部分作了统一的处理,本节将针对2.4节提出的12篇使用DEAP数据库做情绪识别的文献,从特征提取、特征标准化、特征降维、分类器、交叉检验(CV)5个方面进行详细介绍.

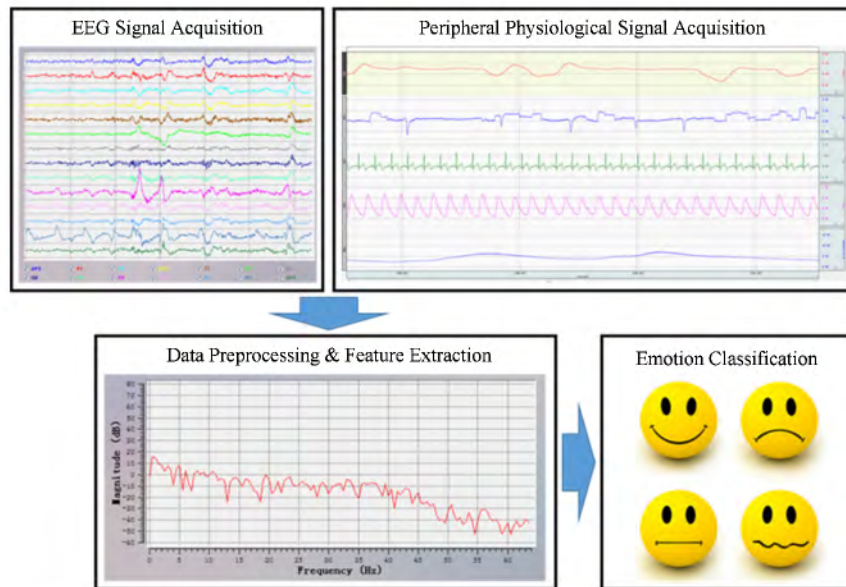


Fig. 2 Main steps of emotion recognition.

图2 情绪识别的主要步骤

3.1 特征提取

特征提取是情绪识别的关键环节,只有找出与诱发情绪相关的敏感、有效特征,才能保证后续情绪识别的准确率.常见的生理信号特征包括时域特征、频域特征和时频域特征.在每一类特征中,我们将依次按照脑电、皮肤电、心电、呼吸、皮肤温度、面部肌电、眼电的顺序进行介绍.

3.1.1 时域特征

脑电信号的时域特征主要集中在事件相关电位(event-related potentials, ERPs)的研究中.ERP研究通常采用单纯的视觉或听觉刺激诱发情绪,并将多个试次的事件进行叠加,分辨率一般为毫秒级.而DEAP数据库用音乐视频诱发情绪,采集的是连续的脑电信号,以频域和时频域特征为主.

外周生理信号的时域特征主要是信号变化的统计特征^[6].常用的皮肤电信号时域特征为信号的均值、中值、标准差、最大最小值、最大最小值的比率和差值等,再将上述信号特征进行一阶差分、二阶差分计算后提取以上相同的统计特征等.心电的时域特征包括心率和心率变异性的均值和标准差.呼吸信号的时域特征主要是通过计算呼吸间隔(即2个相邻的吸气末的波峰的时间差)的均值和标准差.皮肤

温度的时域特征目前在情绪识别领域中应用较少.面部肌电信号通过计算肌电值积分、均方根、过零点、方差、幅值、肌电直方图等指标,反映时间维度上肌电信号振幅的变化特征.眼电信号常用的时域特征包括眨眼频率的均值和标准差.

3.1.2 频域特征

频域特征是由原始生理信号从时域转化而来.在提取脑电频率特征时,一般将脑电信号映射到常用的 $\delta(0.5 \sim 4 \text{ Hz})$, $\theta(4 \sim 8 \text{ Hz})$, $\alpha(8 \sim 12 \text{ Hz})$, $\beta(13 \sim 30 \text{ Hz})$, $\gamma(31 \sim 60 \text{ Hz})$ 等频段上,再从中提取功率、功率谱密度(power spectral density, PSD)、能量及其不对称性(asymmetry, ASM)等特征.皮肤电(galvanic skin response, GSR)信号的频域特征主要是其信号频率范围内(0~2.4 Hz)的谱功率,皮肤导电慢反应(0~0.2 Hz)和极慢反应(0~0.08 Hz)的过零率^[6].心电(electrocardiograph, ECG)的频域分析是将连续正常的R-R间期进行时-频转化,从而获得各频率段上的功率谱密度,如低频段(0~0.04 Hz)、中频段(0.04~0.15 Hz)、高频段(0.15~0.4 Hz)等,用以定量描述心率变异性的能量分布情况.呼吸(respiration, RSP)信号的频域特征主要是其频率范围内(0~2.4 Hz)的谱功率及不同频率段之间能量

差值和比率. 皮肤温度(skin temperature, SKT)的频率特征主要是其频率范围(0~0.2 Hz)内的谱功率. 肌电(electromyogram, EMG)和眼电(electrooculogram, EOG)的频域特征与脑电特征相似, 将原始的信号映射到频域范围内再提取能量、功率、PSD等特征指标.

在基于生理信号的情绪识别研究中, PSD 及其相关指标应用最为广泛. PSD 用来描述其能量特征随频率的变化关系, 与自相关函数互为傅里叶变换的关系. 计算生理信号功率谱的方法可以分为 2 类: 1) 线性估计方法, 包括周期图法、Welch's method、自相关估计、自协方差法等; 2) 非线性估计方法, 如最大似然法、最大熵法等. 线性估计方法是有偏的谱估计方法, 谱分辨率随数据长度的增加而提高; 非线性估计方法大多是无偏的谱估计方法, 可以获得高谱分辨率. 最常用的 PSD 估计方法是使用傅里叶变换技术. 尽管传统的离散傅里叶变换(discrete Fourier transform, DFT)算法能够获取生理信号的频域特征, 但是算法计算量大、耗时长, 不利于计算机对生理信号进行实时处理. 因此, DFT 的一种快速实现算法——快速傅里叶变换(fast Fourier transform, FFT)被发现, 使得 DFT 算法在生理信号的实时分析中得到广泛的应用.

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk}, \quad (1)$$

其中, $k=0, 1, \dots, N-1$; $W_N = e^{-j\frac{2\pi}{N}}$.

式(1)为 DFT 的计算方法, 其中 $x(n)$ 表示输入的离散生理信号序列, W_N 为旋转因子, $X(k)$ 为输入序列 $x(n)$ 对应的 N 个离散频率点的相对幅度. 因为 DFT 计算得到的一组离散频率幅度值实际上是在频率轴上呈周期变化的, 即 $X(k+N) = X(k)$. 因此任意取连续的 N 个点均可以表示 DFT 的计算效果. 一般情况下, 假设 $x(n)$ 来自于低通采样, 采样频率为 f_s , 那么 $X(k)$ 表示在 $[0, f_s - f_s/N]$ 频率上间隔为 f_s/N 的 N 个频率点的相对幅度. 根据式(1)可知, 每计算一个频率点 $X(k)$ 均需要进行 N 次复数乘法和 $N-1$ 次复数加法, 计算 N 个点的 $X(k)$ 共需要 N^2 次复数乘法和 $N(N-1)$ 次复数加法. 当 $x(n)$ 为实数的情况下, 计算 N 点的 DFT 需要 $2 \times N^2$ 次实数乘法, $2N(N-1)$ 次实数加法. 这种算法计算量大而且耗时长.

而 FFT 算法则是利用 W_N 的对称性和周期性(如式(2)所示), 将 N 点的 DFT 变换转为 $\lg N$ 级联的蝶形运算, 每一级均包含有 $N/2$ 次蝶形计算.

由于每一个蝶形运算包含了 1 次复数乘法和 2 次复数加法, 因此计算 N 点 FFT 只需要 $N/2 \times \lg N$ 次复数乘法和 $N \times \lg N$ 次复数加法. 当被采样的序列为实数序列, 只有第一级的计算为实数与复数的混合计算, 经过一次迭代后的计算均变为复数计算, 因此对于输入序列是复数还是实数对 FFT 算法的效率影响较小. 由于计算 N 点的 FFT 只需要 $2N \times \lg N$ 次实数乘法和 $2N \times \lg N$ 次实数加法, 计算效率大大提高.

$$(W_N^{nk})^* = W_N^{-nk},$$

$$W_N^{nk} = W_N^{(n+N)k} = W_N^{n(k+N)}. \quad (2)$$

3.1.3 时频域特征

生理信号大多是非平稳的时变信号, 即不同时刻有不同的频率成分, 因此单纯的时、频分析都不能有效地反映生理信号的这些特征, 而且通过傅里叶变换的时、频域分析方法不可能在时域和频域同时获得较高的分辨率. 越来越多的研究开始将时域和频域联系起来, 对生理信号进行时频特征分析, 着眼于真实信号的时变谱特征. 简单来说就是将一维的时、频域信号表达为二维的时间-频率密度函数, 以揭示生理信号中每个频率分量随时间变化的规律, 如 α 波随时间变化的能量分布. 常用的时频域特征提取方法主要包括: 基于短时傅里叶变换(short-time Fourier transform, STFT)和基于小波包变换(wavelet packet transform, WPT)的时频特征提取方法.

考虑到傅里叶变换处理非平稳信号有天生缺陷, STFT 是在传统傅里叶变换的基础上加入窗函数(即把整个时域过程分解成无数个等长的子过程, 每个子过程近似平稳, 再做傅里叶变换), 通过窗函数的不断移动来决定时变信号局部弦波成分的频率和相位. 由于生理信号的不稳定性, 长时间窗的生理信号变异性大; 而短时间窗又不能提供充足的信息, 因此选取一个合适长度的时间窗对于情绪识别的准确率和计算效率至关重要^[43]. 加窗的方法不仅改善了傅里叶变换处理不稳定信号的缺陷, 也可以用来估计不同情绪状态(如高唤醒度)的起始和持续时间. 特别是当我们使用电影片段或音乐视频诱发情绪时, 不同的刺激材料持续时间不同, 且由于刺激材料的情节不同导致诱发情绪有快有慢. 因此, 通过加窗的方法估算不同情绪状态的起始和持续时间更有其实际意义和应用. 由表 1 可知, 目前使用 DEAP 数据库的情绪识别研究大多选择 1~8 s 的时间窗, 且识别准确率最高的时间窗在 1~2 s^[40-41, 43-44].

基于 STFT 的时频域特征提取的基本思想是: 给一段连续生理信号序列 $x(m)$, 加一个沿时间轴滑动的窗 $w(n-m)$, 窗宽有限, 所得到的傅里叶变换结果 $X(n, \omega)$ 可以表达为

$$X(n, \omega) = \int_{-\infty}^{\infty} w(n-m)x(m)e^{-j\omega m} dm. \quad (3)$$

随着 n 的改变, 窗函数在时间轴上会有位移. 经 $w(n-m)x(m)$ 后, 生理信号只留下了窗函数截取的部分做最后的傅里叶转换. 窗函数种类有很多种, 常见的窗函数有方形、三角形、高斯函数等; 而 STFT 也因窗函数的不同而有不同的名称, 通常没有特别说明的 STFT 为加伯转换 (即窗函数是高斯函数的 STFT).

由于 STFT 的窗口是固定的, 无法满足时变非稳态信号变化的频率需求 (即高频小窗口、低频大窗口). 而小波变换 (wavelet transform, WT) 的窗函数可以随频率的改变而改变, 能对时间 (空间) 频率的局部化分析, 通过伸缩平移运算对生理信号逐步进行多尺度细化, 最终达到高频处时间细分、低频处频率细分, 能自动适应时频信号分析的要求.

WT 算法的基本思想是: 设 $\psi(t)$ 为一平方可积函数, 即 $\psi(t) \in L^2(R)$, 若其傅里叶变换 $\psi(\omega)$ 满足:

$$\int_R \frac{|\psi(\omega)|^2}{|\omega|} d\omega < \infty, \quad (4)$$

则称 $\psi(t)$ 为一个基本小波或小波母函数, 并称式 (4) 是小波函数的可允许条件. 将小波母函数 $\psi(t)$ 进行伸缩和平移, 设其伸缩 (尺度) 因子为 a , 平移因子为 τ , 则平移伸缩后的函数 $\psi_{a,\tau}(t)$ 可表达为

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right), \quad (5)$$

其中, $a, \tau \in R$ 且 $a > 0$; $\psi_{a,\tau}(t)$ 为参数为 a 和 τ 的小波基函数. 由于 a 和 τ 可以取连续变化的值或是限定在离散点上取值, 因此又称之为连续小波变换 (continuous wavelet transform, CWT) 或离散小波变换 (discrete wavelet transform, DWT). 由于小波基具有尺度和位移 2 个参数, 因此将小波基展开意味着将一个时间函数投影到二维的时间-尺度相平面上, 尺度因子 a 对应于频率 (反比), 平移量 τ 对应于时间, 有利于时频域特征的提取.

考虑到 WT 算法在高频段的频率分辨率较差, 而在低频段的时间分辨率较差的缺点, 人们在小波分解的基础上提出了小波包变换 (WPT). 小波包元素是由 3 个参数确定波形, 分别是: 位置、尺度和频率. 对一个给定的正交小波函数, 可以生成一组小波

包基, 每一个小波包基里提供一种特定的信号分析方法, 它可以保存信号的能量并根据特征进行精确的重构. 小波包可以对一个给定的信号进行大量不同的分解. 在正交小波分解过程中, 一般是将低频系数分解为 2 部分: 一个近似系数向量和一个细节系数向量. 在 2 个连续的近似系数中丢失的信息可以在细节系数中得到; 然后将近似系数向量进一步分解为 2 部分, 而细节系数向量不再分解. 在小波包变换中, 每一个细节系数向量也使用近似系数向量分解同样的分法分为 2 部分, 因此它提供了更丰富、更精细的信号分析方法, 提高了信号的时频分辨率.

近年来, 基于双树复小波包变换 (dual-tree complex WPT, DT-CWPT) 也逐步用于提取脑电信号的时频域特征. DT-CWPT 引入了各向异性的小波包分解结构, 由此得到的小波基在各个方向上尺度可以不同, 并具有更多的种类和方向, 能更有效地表现多维度的脑电特征.

3.2 数据标准化

在情绪诱发过程中, 被试通常会在一段时间内观看很多的刺激材料. 以 DEAP 为例, 每名被试需要观看 40 段 1 min 长度的音乐视频, 每两段音乐视频之间允许被试休息 2 min. 也有一些研究会要求被试在休息时间内从事一些不相关的任务, 以尽量让被试脱离上一段材料所诱发的情绪^[37,49]. 尽管如此, 上一段刺激材料诱发的情绪仍然可能会持续到下一段刺激材料的诱发过程中, 尤其是音频、视频材料, 持续时间较长, 情绪卷入可能更大. 此外, 生理信号的个体差异很大, 不同人的生理信号水平各不相同, 甚至同一个人不同时间、不同环境下都会有所不同. 因此, 为了去除之前刺激材料对当前情绪状态的影响以及生理信号的个体差异, 我们通常会将数据进行标准化处理. 最常用的方法是: 用刺激材料呈现时 (即情绪诱发状态下) 的生理信号数据 (如功率谱密度) 减去刺激材料呈现之前 (即平静状态下) 的生理信号数据, 再将差值后的数据统一映射到 $[0, 1]$ 区间上^[46]. 此步骤虽然方法简单, 但意义重大, 可惜大部分的研究并未考虑.

3.3 降维

特征维度过多会导致机器学习建模算法的运行速度变慢, 增加过度拟合的可能性, 导致数据在整个输入空间的分布越稀疏并越难以获得对整个输入空间有代表性的样本. 因此, 通过降维的方法从众多特征维度中选取部分放入建模数据集是必不可少的一步. 常用的降维方法包括线性判别式分析 (linear

discriminant analysis, LDA)、主成分分析(principal component analysis, PCA)、奇异值分解(singular value decomposition, SVD)、列主元 QR 分解(QR factorization with column pivoting, QRcp)等。

3.3.1 线性判别式分析

LDA 也叫做 Fisher 线性判别,是模式识别的经典算法之一。基本思想是:将高维的模式样本投影到最佳鉴别矢量空间,以达到抽取分类信息和压缩特征空间维数的效果,投影后保证模式样本在新的子空间有最大的类间距离 S_b 和最小的类内距离 S_w ,即模式在该空间中有最佳的可分离性。为达到最佳分类效果,LDA 算法要求其所分的类之间耦合度(即类间均值之差的绝对值)低、类内的聚合度(即类内协方差矩阵之和)高。因此,通过引入 Fisher 鉴别准则表达式(式(6)),我们可以将 LDA 算法表达为:选取让 $J_{\text{Fisher}}(\varphi)$ 达到最大值的任一 n 维矢量 φ 作为投影方向,使得投影后的样本具有最大的类间离散度和最小的类内离散度。

$$J_{\text{Fisher}}(\varphi) = \frac{\varphi^T S_b \varphi}{\varphi^T S_w \varphi}. \quad (6)$$

LDA 降维直接和类别的个数相关,与数据本身的维度没有关系。针对多分类问题(c 个类别),LDA 需要由 $c-1$ 个投影向量组成一个投影矩阵,将样本投影到投影矩阵上,从而提取 $c-1$ 维的特征矢量。因此,应用 LDA 降维后的维度的变化范围是 1 到 $c-1$ 。

在 LDA 的假定条件中,每个类内的协方差矩阵是一样的。当我们需要处理更复杂情况时,我们可以考虑每个类内协方差不一样的情况,这时我们需要用到二次判别分析(quadratic discriminate analysis, QDA)算法。QDA 算法在参数估计时需要分别在每一个类中去估计协方差矩阵。

3.3.2 主成分分析

PCA 是另外一种常用的降维算法,与 LDA 算法最大的区别在于:PCA 是一种无监督的映射算法,它所作的只是将整组数据整体映射到最方便表示这组数据的坐标轴上,映射时没有利用任何数据内部的分类信息;而 LDA 是一种有监督的映射算法,在增加了分类信息后,数据之间更易区分。

PCA 算法是根据输入数据的分布给输入数据重新找到更能描述这组数据的正交的坐标轴。其具体思路是:首先求出输入数据的协方差矩阵,再求出这个协方差矩阵的特征值和特征向量。如式(7)所示:假设 λ_i 对应特征值, μ_i 为投影向量。式(7)的左

边取得最大值的条件就是取得最大的特征值 λ_i 。假设要将一个 n 维的数据空间投影到 m 维的数据空间中($m < n$),取前 m 个特征向量构成的投影矩阵就是能够使得数据间方差 S 最大的矩阵。对应最大特征值 λ_i 的那个特征向量 μ_i 的方向就是长轴(也就是主元)的方向,次大特征值的就是第 2 主元的方向,以此类推。

$$S\mu_i = \lambda_i \mu_i. \quad (7)$$

3.4 分类器

实现情绪状态的识别有 3 种方法:1)无监督学习方法。对没有概念标记(分类)的训练样本进行学习,以发现训练样本集中的结构性知识,深度学习网络(deep learning network, DLN)就是典型的无监督学习。DLN 采用与神经网络相似的分层结构,通过逐层特征变换,将样本在原空间的特征表示变换到一个新特征空间,从而使分类或预测更加容易。与人工规则构造特征的方法相比,深度学习利用大数据来学习特征,更能够刻画数据的丰富内在信息^[50]。2)监督学习方法。对具有概念标记(分类)的训练样本进行学习,以尽可能对训练样本集外的数据进行标记(分类)预测。常见的监督学习方法有支持向量机(support vector machine, SVM)、神经网络、决策树、贝叶斯网络、 K -近邻以及隐马尔可夫模型等。3)半监督学习方法。半监督学习方法结合了监督学习与无监督学习的方法,考虑如何利用少量的标注样本和大量的未标注样本进行训练和分类^[51]。

DEAP 数据库是以被试观看完每段音乐视频后对各个维度的评分来标记生理样本,因此基于 DEAP 数据库的情绪分类研究普遍采用监督学习方法,主要包括 SVM、朴素贝叶斯(naïve Bayes, NB)、 K 最近邻(K -nearest neighbors, KNN)、多层感知机(multilayer perceptron, MLP)和人工神经网络(artificial neural network, ANN)等。

3.4.1 K -近邻算法

K -近邻(KNN)算法通过分析最邻近的 k 个样本的类别,决定未知样本的所属类别。如果一个样本在特征空间中的 k 个最相似(即特征空间中最邻近)的样本中的大多数属于某一个类别,则该样本也属于这个类别。具体的思路是:在训练样本中找出最邻近 x_q 的 k 个样本,用 x_1, x_2, \dots, x_k 表示,通过逼近目标函数 $f: \hat{A}^n \rightarrow V \in \{v_1, v_2, \dots, v_n\}$ 计算距离 x_q 最近的 k 个训练样本中最普遍的 f 值(式(8)),以决定待分样本所属的类别。KNN 方法虽然从原理上也依赖于极限定理,但在类别决策时只与极少量的相

邻样本有关. 由于 KNN 方法主要靠周围有限的邻近样本而不是靠判别类域的方法来确定所属类别, 因此对于类域的交叉或重叠较多的待分样本集来说, KNN 方法较其他方法更为适合.

$$\hat{f}(x_q) \leftarrow \arg \max_{v \in V} \delta(v, f(x_i)). \quad (8)$$

模糊最近邻算法 (fuzzy KNN, FKNN) 在 KNN 的基础上考虑各类样本与各类隶属关系的不确定性, 计算近邻样本对各类的隶属度值, 由隶属度函数确定近邻样本与各类的隶属关系. 模糊粗糙近邻算法 (fuzzy rough NN) 在 FKNN 的基础上, 进一步针对各类边缘的粗糙不确定性综合构建模糊粗糙隶属度, 可有效克服分类算法易受离群点和噪声影响的缺陷, 适用于近邻样本隶属关系不确定、各类边缘不清晰的分类问题.

3.4.2 朴素贝叶斯

贝叶斯分类器的分类原理是通过样本的先验概率, 利用贝叶斯公式计算出其后验概率, 即该对象属于某一类的概率, 然后选择具有最大后验概率的类作为未知样本所属的类. 朴素贝叶斯 (NB) 是最简单、应用最广泛的贝叶斯算法, 其前提是假设各属性之间 $x = \{a_1, a_2, \dots, a_n\}$ 互相独立

$$P(y_i | x) = \frac{P(y_i) \prod_{j=1}^n P(a_j | y_i)}{P(x)}. \quad (9)$$

具体的思路是: 首先对于给出的待分类项, 求解在此项出现的条件下各个类别出现的概率 $P(y_1 | x), P(y_2 | x), \dots, P(y_n | x)$, 输出特征属性和训练样本. 计算每个类别在训练样本中的出现频率及每个特征属性划分对每个类别的条件概率估计, 最大概率的类别即为该未知样本的所属类别 $P(y_k | x) = \max\{P(y_1 | x), P(y_2 | x), \dots, P(y_n | x)\}$.

3.4.3 支持向量机

支持向量机 (SVM) 是一种特征空间上间隔最大的线性分类器, 其学习策略便是间隔最大化, 最终可转化为一个凸二次规划问题的求解. 其主要思想可以概括为 2 点: 对于线性可分的情况, 基于结构风险最小化理论, 在特征空间 x_1, x_2, \dots, x_n 中建构最优分割超平面 $f(x) = w^T \cdot x + b$, 使得学习器得到全局最优化, 并且在整个样本空间的期望风险以某个概率满足一定上界, 如式 (10) 所示:

$$\begin{aligned} \max \quad & \frac{1}{\|w\|}, \\ \text{s. t.} \quad & y_i(w^T \cdot x + b) \geq 1, 1 \leq i \leq n. \end{aligned} \quad (10)$$

对于线性不可分的情况, 通过使用非线性映射

算法将低维输入空间线性不可分的样本转化为高维特征空间使其线性可分, 从而使得高维特征空间采用线性算法对样本的非线性特征进行线性分析成为可能. 特征从低维到高维的转换过程中, 我们通常使用核函数 (kernel function) 在低维上进行计算, 而将实质上的分类效果表现在了高维上, 避免了直接在高维空间中的复杂计算.

3.4.4 支持向量回归

除分类问题外, 支持向量也可用来处理回归问题, 即支持向量回归 (SVR) 算法. 与 SVM 相比, SVR 算法的目标也是寻找空间中的最优分割超平面. 不同的是, SVM 寻找的是能将数据分开的平面, 而 SVR 寻找的是能够准确预测数据分布的平面. SVR 算法的主要思想是: 在特征空间 x_1, x_2, \dots, x_n 中建构最优分割超平面 $f(x) = w^T \cdot x + b$, 使其能够准确地预测该特征空间所对应的回归值 y_i , 即 $f(x_i)$ 和 x_i 的差值足够小, 如式 (11) 所示:

$$\begin{aligned} \max \quad & \frac{1}{\|w\|}, \\ \text{s. t.} \quad & \|y_i - (w^T \cdot x + b)\| \leq \epsilon, \end{aligned} \quad (11)$$

其中, $\epsilon \geq 0$, 用来表示 SVR 算法预测值与实际值之间的最大差距. 对于非线性问题, SVR 算法同样可以引用核函数将低维样本升维, 在高维空间中构造线性决策函数来实现非线性回归. 用核函数代替线性方程中的线性项可以使原来的线性算法“非线性化”. 与此同时, 因引进核函数“升维”而增加的可调参数, 其过拟合的风险依然可控.

3.4.5 各种分类器算法的分类效果比较

从表 1 也可看出, SVM 是使用最广泛且识别准确率最高的分类器算法, 其优点有 3 个: 1) 不同于那些力求最小化经验误差的统计学习方法 (例如神经网络和决策树), SVM 能够同时最小化经验误差与最大化几何边缘区. 2) SVM 在处理情绪类别之间数据不平衡方面具有较好的鲁棒性^[39]. 由于无法干预被试的主观评分, 利用被试评分标定训练样本的过程很可能会导致情绪类别之间的数据不平衡. 以 DEAP 数据库为例, 其高、低唤醒度的数据比例为 63.9% 和 36.1%, 正负特性 (效价) 的数据比例为 63.1% 和 36.9%. 由于 SVM 的最终决策函数只由少数的支持向量所确定, 其支持向量的样本集具有一定的鲁棒性. 3) SVM 在基于生理大数据的情绪实时识别应用中优势明显. 实时的情绪识别要求在短时间内从大规模生理信号中提取干净、真实、可信的特征并以

此训练、建立分类模型并预测情绪状态. 实时决定了训练样本有限, SVM 不需要耗费太多的运算时间求解支持向量, 而 SVM 少数支持向量(而不是样本空间的维数)决定最终结果的特性有效地避免了生理大数据维数高的缺陷, 帮助我们剔除冗余样本的同时抓住关键样本.

3.5 交叉检验

交叉验证(CV)的目的是为了得到稳定可靠的模型. N 折交叉验证(N -fold CV)是比较常用的交叉检验方法. 它将初始采样分割成 N 个子样本, 一个单独的子样本被保留作为验证模型的数据, 其他 $N-1$ 个样本用来训练. 交叉验证重复 N 次, 每个子样本验证一次, 平均 N 次的结果或者使用其它结合方式, 最终得到一个单一估计. 被试依赖(subject dependent, SD)是指留一个试次作为测试集, 其他的试次作为训练集, 然后重复这个过程, 直至所有的试次都做了一次测试集, 正确率为所有试次的平均. 被试独立(subject independent, SID)留一个被试作为测试集, 其他的被试作为训练集, 然后重复这个过程, 直至所有的被试都做了一次测试集, 正确率为所有被试的平均. 一般来讲, 被试依赖的正确率较高, 但推广性差. 而实时在线分析的正确率低于线下分析的正确率.

4 应 用

随着基于生理信号的情绪识别研究逐渐深入, 越来越多的研究者们开始思考如何将研究成果转化和应用. 传统采集生理信号的方式是在实验室环境下完成的, 采集需要一系列的准备过程和繁琐的程序, 花费时间较长, 收集到的生理信号多数用于线下分析. 近年来, 微针型干电极技术的不断发展以及可穿戴设备的与时俱进, 使得情绪识别的相关研究成果更多地应用到实际产品中. 干电极生理信号采集系统采用无线通信技术, 使用简便, 不易受环境制约. Jatupaiboon 等人^[52]采用 14 导干电极脑电帽(emotiv)采集被试的脑电数据, 建立了实时监测高兴-不高兴(即正性-负性)识别系统. 研究者基于这个系统开发了 2 款游戏: 头像游戏和奔跑游戏. 头像游戏是用户高兴时呈现高兴的面孔和音乐, 不高兴时则呈现不高兴的面孔和音乐; 奔跑游戏则是用户越高兴, 则游戏角色跑的越快. 值得说明的是, emotiv 设备本身也能够实时地识别出多种情绪状态.

尽管干电极技术和可穿戴设备目前还有很多局限, 如信号稳定性不强、电源供电时间较短等, 限制了其应用范围^[53]; 但未来随着技术的不断革新, 基于生理信号的情绪识别必将在社交网络、教育、医疗、心理治疗、军事等领域发挥重要作用.

5 讨论与展望

从已有文献来看, 基于生理信号的情绪识别研究还处于摸索阶段, 研究很多但良莠不齐. 由于情绪诱发材料、采集设备、生理信号的特征提取、降维和分类算法等方面的差异性, 已有研究中的情绪识别正确率高低不一, 无法进行横向比较. 因此, 本文建议未来的情绪研究可以从以下 2 个方面展开:

1) 建议研究人员使用标准化的刺激材料, 特别是生态效度更高的电影片段和音乐视频. 未来预期还要建立适用于中国社会和文化背景的华语电影片段和音乐视频标准库.

2) 建议研究人员使用公开发表的生理信号数据库(如 DEAP 数据库)进行特征提取算法、分类算法和模型的开发, 既能避免情绪诱发和信号采集过程中的差异性, 也方便研究人员对各个算法模型进行横向比较. 未来预期能够建立基于华语电影片段和音乐视频标准刺激的生理信号库, 提供包括脑电和外周生理信号在内的生理大数据库.

参 考 文 献

- [1] Minsky M. Society of Mind [M]. New York: Simon and Schuster, 1988
- [2] Picard R W. Affective Computing [M]. Cambridge, MA: MIT Press, 1997
- [3] Kim M K, Kim M, Oh E. et al. A review on the computational methods for emotional state estimation from the human EEG [J]. Computational and Mathematical Methods in Medicine, 2013, 2013: 1-13
- [4] Nie Dan, Wang Xiaowei, Duan Ruonan, et al. A survey on EEG based emotion recognition [J]. Chinese Journal of Biomedical Engineering, 2012, 31(4): 595-606 (in Chinese) (聂聃, 王晓韩, 段若男, 等. 基于脑电的情绪识别研究综述 [J]. 中国生物医学工程学报, 2012, 31(4), 595-606)
- [5] Zhang Di, Wan Baikun, Ming Dong. Research progress on emotion recognition based on physiological signals [J]. Journal of Biomedical Engineering, 2015, 32(1): 229-234 (in Chinese)

- (张迪, 万柏坤, 明东. 基于生理信号的情绪识别研究进展[J]. 生物医学工程学杂志, 2015, 32(1): 229-234)
- [6] Koelstra S, Muhl C, Soleymani M, et al. DEAP: A database for emotion analysis using physiological signals [J]. IEEE Trans on Affective Computing, 2012, 3(1): 18-31
- [7] Gupta R, Gupta H, Mohania M. Cloud computing and big data analytics: What is new from databases perspective? [C] //Proc of Big Data Analytics. Berlin: Springer, 2012: 42-61
- [8] Izard C E. The many meanings/aspects of emotion: Definitions, functions, activation, and regulation [J]. Emotion Review, 2010, 2(4): 363-370
- [9] van den Broek E L. Ubiquitous emotion-aware computing [J]. Personal and Ubiquitous Computing, 2013, 17(1): 53-67
- [10] Mehrabian A. Pleasure arousal dominance: A general framework for describing and measuring individual differences in temperament [J]. Current Psychology, 1996, 14(4): 261-292
- [11] Liu Ye, Fu Qiufang, Fu Xiaolan. The interaction between cognition and emotion [J]. Chinese Science Bulletin, 2009, 54(18): 2783-2796 (in Chinese)
(刘烨, 付秋芳, 傅小兰. 认知与情绪的交互作用[J]. 科学通报, 2009, 54(18): 2783-2796)
- [12] Collet C, Vernet-Maury E, Delhomme G, et al. Autonomic nervous system response patterns specificity to basic emotions [J]. Journal of the Autonomic Nervous System, 1997, 62(1/2): 45-57
- [13] Barrett L F. Are emotions natural kinds? [J]. Perspectives on Psychological Science, 2006, 1(1): 28-58
- [14] Liu Yisi, Sourina O, Nguyen M K. Real-time EEG-based human emotion recognition and visualization [C] //Proc of Int Conf on Cyberworlds. Piscataway, NJ: IEEE, 2010: 262-269
- [15] Fontaine J R J, Scherer K R, Roesch E B, et al. The world of emotions is not two-dimensional [J]. Psychological Science, 2007, 18(12): 1050-1057
- [16] Yin Dezhi, Bond Samuel D, Zhang Han. Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews [J]. Mis Quarterly, 2014, 38(2): 539-560
- [17] Mauss I B, Robinson M D. Measures of emotion: A review [J]. Cognition & Emotion, 2009, 23(2): 209-237
- [18] Hamann S. Mapping discrete and dimensional emotions onto the brain: Controversies and consensus [J]. Trends in Cognitive Sciences, 2012, 16(9): 458-466
- [19] AlZoubi O, Calvo R A, Stevens R H. Classification of EEG for affect recognition: An adaptive approach [C] //Proc of AI 2009: Advances in Artificial Intelligence. Berlin: Springer, 2009: 52-61
- [20] Chanel G, Kierkels J J M, Soleymani M, et al. Short-term emotion assessment in a recall paradigm [J]. International Journal of Human-Computer Studies, 2009, 67(8): 607-627
- [21] Kothe C A, Makeig S, Onton J A. Emotion recognition from EEG during self-paced emotional imagery [C] //Proc of Humaine Association Conf on Affective Computing and Intelligent Interaction. Piscataway, NJ: IEEE, 2013: 855-858
- [22] van Reekum C, Johnstone T, Banse R, et al. Psychophysiological responses to appraisal dimensions in a computer game [J]. Cognition and Emotion, 2004, 18(5): 663-688
- [23] Merckx P A B, Truong K P, Neerinx M A. Inducing and measuring emotion through a multiplayer first-person shooter computer game [C] //Proc of Computer Games Workshop. Amsterdam, the Netherlands: TNO Defensie en Veiligheid, 2007
- [24] Lang P J, Bradley M M, Cuthbert B N. International affective picture system (IAPS): Technical manual and affective ratings [R]. Gainesville, FL: NIMH Center for the Study of Emotion and Attention, 1997
- [25] Lang P J, Bradley M M, Cuthbert B N. International affective picture system (IAPS): Affective ratings of pictures and instruction manual [R]. Gainesville, FL: NIMH Center for the Study of Emotion and Attention, 2005
- [26] Ekman P, Friesen W. Pictures of Facial Affect [M]. Sunnyvale, CA: Consulting Psychologists Press, 1975
- [27] Bradley M M, Lang P J. Affective norms for English words (ANEW): Instruction manual and affective ratings, C-1 [R]. Gainesville, FL: the Center for Research in Psychophysiology, University of Florida, 1999
- [28] Ekman P. Emotions Revealed: Understanding Thoughts and Feelings [M]. London: Weidenfeld and Nicolson, 2003
- [29] Bai Lu, Ma Hui, Huang Yuxia, et al. The development of native Chinese affective picture system—a pretest in 46 college students [J]. Chinese Mental Health Journal, 2005, 19(11): 719-722 (in Chinese)
(白露, 马慧, 黄宇霞, 等. 中国情绪图片系统的编制——在46名中国大学生中的试用[J]. 中国心理卫生杂志, 2005, 19(11): 719-722)
- [30] Zheng Pu, Liu Conghui, Yu Guoliang. An overview of mood-induction methods [J]. Advances in Psychological Science, 2012, 20(1): 45-55 (in Chinese)
(郑璞, 刘聪慧, 俞国良. 情绪诱发方法述评. 心理科学进展, 2012, 20(1): 45-55)
- [31] Bradley M M, Lang P J. The International affective digitized sounds (2nd ed; IADS-2): Affective ratings of sounds and instruction manual, B-3 [R]. Gainesville, FL: University of Florida, 2007

- [32] Liu Taosheng, Ma Hui, Huang Yuxia, et al. Primary study of establishing an affective sound system of China [J]. Chinese Mental Health Journal, 2006, 20(11): 709-712 (in Chinese)
(刘涛生, 马慧, 黄宇霞, 等. 建立情绪声音刺激库的初步研究[J]. 中国心理卫生杂志, 2006, 20(11): 709-712)
- [33] Khosrowabadi R, Quek H C, Wahab A, et al. EEG-based emotion recognition using self-organizing map for boundary detection [C] //Proc of Int Conf on Pattern Recognition. Piscataway, NJ: IEEE, 2010: 4242-4245
- [34] Gross J J, Levenson R W. Emotion elicitation using films [J]. Cognition & Emotion, 1995, 9(1): 87-108
- [35] Schaefer A, Nils F, Sanchez X, et al. Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers [J]. Cognition and Emotion, 2010, 24(7): 1153-1172
- [36] Xu Pengfei, Huang Yuxia, Luo Yuejia. Establishment and assessment of native Chinese affective video system [J]. Chinese Mental Health Journal, 2010, 24(7): 551-561 (in Chinese)
(徐鹏飞, 黄宇霞, 罗跃嘉. 中国情绪影像材料库的初步编制和评定[J]. 中国心理卫生杂志, 2010, 24(7): 551-561)
- [37] Liang Yuchi, Hsieh Shulan, Weng Chiaying, et al. Taiwan corpora of Chinese emotions and relevant psychophysiological data—standard Chinese emotional film clips database and subjective evaluation normative data [J]. Chinese Journal of Psychology, 2013, 55(4): 601-621 (in Chinese)
(梁育绮, 谢淑兰, 翁嘉英, 等. 台湾地区华人情绪与相关心理生理数据库—标准化华语版情绪电影短片材料库与主观评量常模[J]. 中华心理期刊, 2013, 55(4): 601-621)
- [38] Daimi S N, Saha G. Recognition of emotions induced by music videos using DT-CWPT [C] //Proc of Indian Conf on Medical Informatics and Telemedicine. Piscataway, NJ: IEEE, 2013: 53-57
- [39] Daimi S N, Saha G. Classification of emotions induced by music videos and correlation with participants' rating [J]. Expert Systems with Applications, 2014, 41(13): 6057-6065
- [40] Rozgic V, Vitaladevuni S N, Prasad R. Robust EEG emotion classification using segment level decision fusion [C] //Proc of IEEE Int Conf on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE, 2013: 1286-1290
- [41] Zhuang Xiaodan, Rozgic V, Crystal M. Compact unsupervised EEG response representation for emotion recognition [C] //Proc of IEEE EMBS Int Conf on Biomedical and Health Informatics. Piscataway, NJ: IEEE, 2014: 736-739
- [42] Torres C A, Orozco A A, Alvarez M A. Feature selection for multimodal emotion recognition in the arousal-valence space [C] //Proc of Int Conf of the IEEE Engineering in Medicine and Biology Society. Piscataway, NJ: IEEE, 2013: 4330-4333
- [43] Yoon H J, Chung S Y. EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm [J]. Computers in Biology and Medicine, 2013, 43(12): 2230-2237
- [44] Hatamikia S, Maghooli K, Nasrabadi A M. The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals [J]. Journal of Medical Signals and Sensors, 2014, 4(3): 194-201
- [45] Xiang Jie, Cao Rui, Li Li. Emotion recognition based on the sample entropy of EEG [J]. Bio-Medical Materials and Engineering, 2014, 24(1): 1185-1192
- [46] Jirayucharoensak S, Pan-Ngum S, Israsena P. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation [J]. The Scientific World Journal, 2014, 2014: 1-10
- [47] Matiko J W, Beeby S P, Tudor J. Fuzzy logic based emotion classification [C] //Proc of IEEE Int Conf on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE, 2014: 4389-4393
- [48] Verma G K, Tiwary U S. Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals [J]. Neuroimage, 2014, 102: 162-172
- [49] Lee Youyun, Hsieh Shulan. Classifying different emotional states by means of EEG-based functional connectivity patterns [J]. Plos One, 2014, 9(4): e95415
- [50] Yu Kai, Jia Lei, Chen Yuqiang, et al. Deep learning: Yesterday, today, and tomorrow [J]. Journal of Computer Research and Development, 2013, 50(9): 1799-1804 (in Chinese)
(余凯, 贾磊, 陈雨强, 等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展, 2013, 50(9): 1799-1804)
- [51] Li Ming, Zhou Zhihua. Online semi-supervised learning with multi-kernel ensemble [J]. Journal of Computer Research and Development, 2008, 45(12): 2060-2068 (in Chinese)
(黎铭, 周志华. 基于多核集成的在线半监督学习方法[J]. 计算机研究与发展, 2008, 45(12): 2060-2068)
- [52] Jatupaiboon N, Pan-Ngum S, Israsena P. Real-time EEG-based happiness detection system [J]. The Scientific World Journal, 2013, 2013: 1-10
- [53] Wang Xiaowei, Shi Lichen, Lu Baoliang. A survey on the technology of dry electrodes for EEG recording [J]. Chinese Journal of Biomedical Engineering, 2010, 29(5): 777-784 (in Chinese)

(王晓韩, 石立臣, 吕宝粮. 干电极脑电采集技术综述[J]. 中国生物医学工程学报, 2010, 29(5): 777-784)



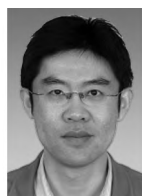
Zhao Guozhen, born in 1984. Received his PhD degree from the State University of New York in 2011. Assistant Professor at the Institute of Psychology, Chinese Academy of Sciences. His main research interests include the mathematical modeling of human cognition and performance, transportation safety, human computer interaction, neuroergonomics and their applications in intelligent system design.



Song Jinjing, born in 1990. Master candidate at the Institute of Psychology, Chinese Academy of Sciences. Her main research interests include emotion recognition based on EEG and physiological signals.



Ge Yan, born in 1981. Received her PhD degree from the Institute of Psychology, Chinese Academy of Sciences in 2008. Associate professor at Institute of Psychology, Chinese Academy of Sciences. Her main research interests include human-computer interaction.



Liu Yongjin, born in 1977. Received his PhD degree from the Hong Kong University of Science and Technology in 2004. Associate professor at the Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Computer Science and Technology, Tsinghua University. Member of IEEE and IEEE Computer Society. His main research interests include computational geometry, cognitive computation and computer graphics.



Yao Lin, born in 1985. Received his PhD degree from the Institute of Psychology, Chinese Academy of Sciences in 2012. Researcher at China Mobile Research Institute. His main research interests include human-computer interaction and innovation in handheld, mobile and wearable technologies.



Wen Tao, born in 1985. Received his master degree at the Capital Normal University in 2010. Researcher at China Mobile Research Institute. His main research interests include collecting, analyzing, presenting usability metrics of mobile data services through behavioral-based research method.